AD-A120 381 ILLINOIS UNIV AT CHICAGO CIRCLE DEPT OF MATHEMATICS F/G 9/4
ON THE SUPERADDITIVITY OF INFORMATION MATRICES IN GAUSS-MARKOV --ETC(U)
1982 A S HEDAYAT, D MAJUMDAR AFOSR-80-0170
UNCLASSIFIED AFOSR-TR-82-0891 NL

1 of 1

END
DATE
FILMED
11 82

ON THE SUPERADDITIVITY OF INFORMATION MATRICES IN GAUSS-MARKOV MODELS.[1]

(4)

by

A.S. Hedayat and Dibyen Majumdar

Department of Mathematics, Statistics
and Computer Science
University of Illinois, Chicago

ABSTRACT. The statistical efficiency of designs associated with the linear model is by and large measured solely on the basis of the information matrices. It is shown that when data from two experiments with the same model, which might contain nuisance parameters, but with possibly different design matrices are combined, then the resulting information matrix is larger (in the sense of nonnegative definiteness) than the sum of the individual information matrices. Cases where equality is achieved are completely characterized geometrically and statistically. Conditions where the best linear unbiased estimator of estimable functions are obtained as linear combination of the best linear unbiased estimators of the same function from the individual experiments are determined.

Key words and phrases: Classical Gauss-Markov Model. Nuisance Parameters. Information Matrix. Estimable Functions. BLUE. Combining Experiments. Confounding.

1 This research was supported by Grant No. AFOSR 80-0170.

- A -

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 82-0891	2. GOVT ACCESSION NO. AD-A120381	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON THE SUPERADDITIVITY OF INFORMATION MATRICES IN GAUSS-MARKOV MODELS		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A.S. Hedayat and Dibyen Majumdar		8. CONTRACT OR GRANT NUMBER(s) AFOSR-80-0170
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics, Statistics, and Computer Science, University of Illinois, Box 4348, Chicago IL 60680		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Directorate of Mathematical & Information Sciences Air Force Office of Scientific Research Bolling AFB DC 20332		12. REPORT DATE . 1982
		13. NUMBER OF PAGES 44
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION ST. 4ENT (of abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY TES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Classical Gauss-Markov model; nuisance parameters; information matrix; estimable functions; BLUE; combining experiments; confounding.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The statistical efficiency of designs associated with the linear model is by and large measured solely on the basis of the information matrices. It is shown that when data from two experiments with the same model, which might contain nuisance parameters, but with possibly different design matrices are combined, then the resulting information matrix is larger (in the sense of nonnegative definiteness) than the sum of the individual information matrices. Cases where equality is achieved are completely characterized geometrically and statistically. Conditions where the best linear unbiased estimator of (CONTINUED)

DD FORM 1 JAN 73 1473

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ITEM #20, CONTINUED: estimable functions are obtained as linear combination of the best linear unbiased estimators of the same function from the individual experiments are determined.

ON THE SUPPERADDITIVITY OF INFORMATION MATRICES IN GAUSS-MARKOV MODELS

By

A.S. Hedayat and Dibyen Majumdar

Department of Mathematics, Statistics
and Computer Science
University of Illinois, Chicago

1. INTRODUCTION

Scientists perform experiments to gather information about some underlying phenomenon. What constitutes information largely depends on the problem at hand and the mode of inference to be made. However, one thing is obvious that, no matter what the definition of information, it should be nondecreasing in the number of observations. Here, we shall exclude from our consideration cost and any other nonstatistical/mathematical matter. For example, the notion "amount of information per $" is of no consideration to us. With this philosophy scientists working on identical problems should combine their data for the purpose of analysis and inference.

There are various notions of information which have found popularity amoung scientists. Before adopting any such notion one has to establish the minimum requirement that it is nondecreasing in the number of observations. Statisticians dealing with experiments whose data follow the linear model, by and large, have adopted the inverse of the variance as the amount of information. Or more generally, the "smaller" the variance-covariance matrix the more information has been obtained. Equivalently, the "bigger"

the information matrix (the inverse of the variance-covariance matrix) the more information is provided. Here, we have adopted the following concept. If A and B are two nonnegative definite matrices then we say A is at least as large as B (written as $A \geq B$) if $A-B$ is a nonnegative definite matrix. A is bigger than B $(A>B)$ if $A-B$ is positive definite.

As we noted earlier if D_1 and D_2 are two sets of data for the same problem they should be combined for the purpose of analysis and inference since information, $\mathcal{I}$, provided by $D_1 \cup D_2$ is as large as $\mathcal{I}_1$ or $\mathcal{I}_2$, the information provided by D_1 and D_2 respectively. Thus in general $\mathcal{I} \geq \mathcal{I}_i, i = 1,2$. In view of the above a natural question is this. Are there settings for which one can relate $\mathcal{I}$ to $\mathcal{I}_1 + \mathcal{I}_2$. This paper answers this question in affirmative. We establish that for the classical linear model there is a <u>supperadditivity</u> for the information matrix of the combined data, i.e., $\mathcal{I} \geq \mathcal{I}_1 + \mathcal{I}_2$ whether or not there are nuisance parameters. We prove this in Section 3. In this section we also give both statistical and geometrical characterizations of cases in which $\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2$.

In Section 4 we study the status of Gauss-Markov theorem for the combined data and shall explore its relation to associated Gauss-Markov theorems for the individual sets of data. Explicity, we provide answers to the following questions. (i) Given <u>a</u> linear parametric function $p'\theta_1$ (θ_2 is the vector of nuisance parameters) which is estimable under both sets of data, then under

what conditions can the BLUE of $p'\theta_1$ under the combined data be expressed as a linear combination of BLUES of $p'\theta_1$ under individual sets of data. (ii) The same question as in (i) except that we insist the same conclusion be true for all estimable functions. Besides characterization of involved cases we give explicit forms of the linear functions of BLUEs so that we can save some computational time in case for each set of data the BLUEs have already been computed.

In Section 5 we indicate how the results in Sections 3 and 4 could be utilized in the area of optimal design of experiments. We give examples to illustrate our results throughout the paper. To develop our theory we need some preliminary results which we have summarized in Section 1.

2. PRELIMINARY RESULTS.

In this section we shall introduce the model, some notations and state some known and not so well known results, which are needed in the subsequent sections.

We shall be using the classical Gauss-Markov model throughout this paper. This may be described as follows:

$$Y = X_1\theta_1 + X_2\theta_2 + \epsilon \ . \tag{2.1}$$

Here Y is the vector of observations. X_1 is the design matrix associated with θ_1, which is a vector of all parameters of

interest. X_2 is the design matrix associated with θ_2, which is a vector of nuisance or covariate parameters. ϵ is the usual error vector with $E(\epsilon) = 0$, which we assume to be homoscedastic, i.e., $V(\epsilon) = \sigma^2 I$, where I dentes the identity matrix whose dimension will be clear from the context.

Throughout we adopt the following notation: For a matrix A, $\mathcal{M}(A)$ will denote the column space of A, $r(A)$ the rank of A and A^- a generalized inverse of A. The matrix P_A will denote the orthogonal projection operator onto $\mathcal{M}(A)$. More explicitly, $P_A = A(A'A)^- A'$. Orthogonality for us, is always in terms of the dot product. If A, B and C are square matrices then the block diangonal matrix

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

will be denoted by diag (A,B). Diag (A,B,C) will have a similar meaning.

We shall now state and prove four lemmas. Lemmas 2.1 and 2.2 describe well known results of Gauss-Markov models in a form suitable for us. Lemmas 2.3 and 2.4 are purely algebraic in nature.

<u>Lemma 2.1</u>. $b'Y$ is the BLUE of the linear parametric function $p'\theta$ for the Gauss-Markov model

$$Y = X\theta + \epsilon \ , \quad E(\epsilon) = 0 \ , \quad V(\ ) = \sigma^2 I$$

iff $$X'b = p \quad \text{and} \quad b \in \mathcal{M}(X).$$

Proof. $b'Y$ is unbiased for $p'\theta$ iff $X'b = p$. Then, by Lehmann-Scheffe Theorem, $b'Y$ is BLUE of $p'\theta$ iff $COV(b'Y, z'Y) = 0$ for every z for which $E(z'Y) = 0$ for all θ. i.e., $b'z = 0$ for every z such that $z'X = 0$. Note that the condition of the Lemma on b satisfies the Lehmann-Scheffe condition since if $b \in \mathcal{M}(X)$, then obviously $b'z = 0$, whenever $z'X = 0$. Conversely, given b write $b = b_1 + b_2$, where $b_1 \in \mathcal{M}(X)$ and $b_2'X = 0$. Hence by Lehmann-Scheffe theorem

$$0 = b'b_2 = b_2'b_2 ,$$

which implies $b_2 = 0$, i.e., $b = b_1 \in \mathcal{M}(X)$.

Lemma 2.2. (i) $b'Y$ is BLUE of the parametric function $p'\theta_1$ for the Gauss-Markov model (2.1) iff

$$X_1'b = p \; ; \; X_2'b = 0 \quad \text{and} \quad b \in \mathcal{M}(X_1 : X_2)$$

(ii) $b'Y$ is BLUE of $E(b'Y)$ for the model (2.1), iff

$$b \in \mathcal{M}[(I - P_{X_2})X_1]$$

Proof. (i) Unbiasedness of $b'Y$ implies

$$b'X_1\theta_1 + b'X_2\theta_2 = p'\theta_1 , \quad \text{for all } \theta_1, \theta_2$$

Hence $X_1'b = p$ and $X_2'b = 0$. The rest of (i) follows from Lemma 2.1. Observe that $X_2'b = 0$ iff $b = (I-P_{X_2})b$. But $b = X_1h_1 + X_2h_2$ for some h_1, h_2, since $b \in \mathcal{M}(X_1:X_2)$. Hence $b = (I-P_{X_2})b = (I-P_{X_2})X_1h_1 \in \mathcal{M}[(I-P_{X_2})X_1]$. Conversely if $b'Y = h'X_1'(I-P_{X_2})Y$, then Lehmann-Scheffe Theorem can be used exactly as in Lemma 2.1 to show that $b'Y$ is BLUE of $E(b'Y)$.

Lemma 2.3. Let A and B be two matrices such that $\mathcal{M}(A') = \mathcal{M}(B')$. Then for each $p \in \mathcal{M}(A')$, there exists a vector h and a scalar α (depending on p) such that

$$\begin{aligned} A'A\,h &= \alpha p \\ B'B\,h &= (1-\alpha)p \end{aligned} \tag{2.2}$$

are simultaneously satisfied, iff $B'B = \lambda\, A'A$ for some $\lambda > 0$.

Proof. First note that $\mathcal{M}(A') = \mathcal{M}(A'A)$ and $\mathcal{M}(B') = \mathcal{M}(B'B)$, so $p \in \mathcal{M}(A'A) = \mathcal{M}(B'B)$. Also, if (2.2) is satisfied for some h, for a given non null p, then $p'h = \alpha p'(A'A)^{-}p = (1-\alpha)p'(B'B)^{-}p$. Solving for α, we obtain

$$\alpha = \frac{p'(B'B)^{-}p}{p'(A'A)^{-}p + p'(B'B)^{-}p} \tag{2.3}$$

Hence (2.2) implies that α is of the form (2.3). We shall now prove the 'only if' part of the theorem. It is well known that

any two nonnegative definite matrices can be simultaneosly diagonalized by a single nonsingular matrix (see Rao and Mitra (1971), p. 122). So, let T be a nonsingular matrix such that

$$A'A = T' \operatorname{diag}(D_1,0)T \ , \ B'B = T' \operatorname{diag}(D_2,0)T \qquad (2.4)$$

for positive definite diagonal matrices D_1 and D_2 of the same dimension (say r), since $\mathcal{M}(A'A) = \mathcal{M}(B'B)$. Equation (2.2) may be written as

$$\begin{aligned} T' \operatorname{diag}(D_1,0)\ T\ h &= \alpha p \\ T' \operatorname{diag}(D_2,0)\ T\ h &= (1-\alpha)p \end{aligned} \qquad (2.5)$$

Suppose $A'A$ and $B'B$ are $m \times m$ matrices. Partition

$$T' = (T_1' : T_2')$$

where T_1' is $m\times r$ and T_2' is $m\times m-r$. Let $s = T h$, $q = (T')^{-1}p = Qp$, where $Q^{-1} = T'$. Corresponding to the partition of T, partition $s' = (s_1' : s_2')$, $q' = (q_1' : q_2')$, $Q' = (Q_1' : Q_2')$. Then $QT' = I$ implies $Q_1T_1' = I$, $Q_2T_1' = 0$. Since by (2.4) $\mathcal{M}(A') = \mathcal{M}(T_1')$, $p = T_1'w$ for some w. Hence $q_2 = Q_2p = 0$ and $q_1 = Q_1p = w$. So as p varies over the whole of $\mathcal{M}(T_1')$, q_1 varies over the entire euclidean space R^r. Equation (2.5) reduces to

$$D_1\, s_1 = \alpha\, q_1$$
$$D_2\, s_1 = (1-\alpha)\, q_1 \tag{2.6}$$

If there exists h, α such that (2.2) is satified for all $p \in \mathcal{M}(A')$, then there exists s_1, α such that (2.6) is satisfied for all $q_1 \in R^r$. Given $q_1 \neq 0$, it follows from (2.6) that

$$s_1 = \alpha\, D_1^{-1}\, q_1 = (1-\alpha)\, D_2^{-1}\, q_1$$

i.e.,
$$D_2 D_1^{-1}\, q_1 = \alpha^{-1}(1-\alpha) q_1 \tag{2.7}$$

Note that $\alpha^{-1}(1-\alpha) > 0$, since from (2.3) $0 < \alpha < 1$. Write $D_j = \mathrm{diag}(d_{j1}, \ldots, d_{jr})$, $j = 1,2$. Since (2.7) is satisfied for all $q_1 \in R^r$, $d_{2i}/d_{1i} = \alpha^{-1}(1-\alpha)$, for all $i = 1, \ldots, r$, by choosing $q_1' = (1,1,\ldots,1)$, a vector of one's. Hence $D_2 = \alpha^{-1}(1-\alpha) D_1$, which implies $B'B = \lambda\, A'A$, with $\lambda = \alpha^{-1}(1-\alpha)$.

To prove the if part, notice that if $B'B = \lambda\, A'A$ for some $\lambda > 0$, then the only admissible value of α, by equation (2.3), is

$$\alpha = \lambda^{-1}(1 + \lambda^{-1})^{-1} = (1 + \lambda)^{-1} \tag{2.8}$$

If h is chosen to satisfy $A'A\, h = \alpha p$, then $B'B\, h = \lambda\, A'A\, h = \lambda\, \alpha\, p = (1-\alpha)p$, by equation (2.8). Hence (2.2) is satisfied whatever $p \in \mathcal{M}(A')$.

Let A and B be two matrices with the same number of columns. Then the following is well known (Rao(1965) p. 34):

$$\mathcal{M}(A') \cap \mathcal{M}(B') = \{0\} \Rightarrow A'A(A'A + B'B)^{-} A'A = A'A .$$

In the following Lemma the converse is also established.

Lemma 2.4. The following are equivalent.

(i) $\mathcal{M}(A') \cap \mathcal{M}(B') = \{0\}$

(ii) $A'A(A'A + B'B)^{-}A'A = A'A$ for some g-inverse of $A'A + B'B$.

Proof. To prove (ii) $\Rightarrow$ (i) observe that if (ii) is satisfied for some g-inverse of $A'A + B'B$, then it is satisfied for all g-inverses. Let T be a nonsingular matrix such that $A'A = T' \operatorname{diag}(D_1,0,0)T$, $B'B = T' \operatorname{diag}(D_2,D_3,0)T$ where D_1,D_2,D_3 are diagonal matrices. D_1 and D_2 are of the same dimension, D_1 is positive definite. Then $A'A + B'B = T' \operatorname{diag}(D_1 + D_2,D_3,0)T$. Choose the following g-inverse

$$(A'A + B'B)^{-} = T^{-1} \operatorname{diag}((D_1+D_2)^{-1}, D_3^{-}, 0)T'^{-1} .$$

Then (ii) implies

$$\operatorname{diag}(D_1,0,0) \operatorname{diag}((D_1+D_2)^{-1},D_3^{-},0) \operatorname{diag}(D_1,0,0) = \operatorname{diag}(D_1,0,0)$$

i.e., $$D_1(D_1+D_2)^{-1} D_1 = D_1$$

i.e., $$D_2 = 0$$

Hence if $T' = (T_1' : T_2' : T_3')$ in the partitioned form, then

$$A'A = T_1' D_1 T_1 \quad \text{and} \quad B'B = T_2' D_3 T_2 .$$

Thus $\mathcal{M}(A'A) = \mathcal{M}(T_1')$ and $\mathcal{M}(B'B) \subset \mathcal{M}(T_2')$. Hence (i) follows.

3. A LOWER BOUND ON INFORMATION MATRICES

The main purpose of this section is to prove that the concept of information in the context of linear models defined in the introduction is superadditive, i.e., the information associated with the combination of two experiments is at least equal to the sum of the information provided by individual experiments. We also examine the statistical and geometrical interpretation of those cases where equality is achieved. Various examples are provided to elucidate the theory.

Let there be two experiments with the following models:

Experiment 1: $Y_1 = X_{11}\theta_1 + X_{12}\theta_2 + \epsilon_1 , E(\epsilon_1) = 0 , V(\epsilon_1) = \sigma^2 I$

Experiment 2: $Y_2 = X_{21}\theta_1 + X_{22}\theta_2 + \epsilon_2 , E(\epsilon_2) = 0 , V(\epsilon_2) = \sigma^2 I$

The meaning of the symbols in each model is the same as in equation (2.1). The number of observations in experiment 1 is n_1 and in experiment 2 is n_2. It is well known that the information

matrices for estimable functions of θ_1 for the first and second experiments are respectively

$$\mathscr{I}_1 = X_{11}'(I_{n_1} - X_{12}(X_{12}'X_{12})^{-}X_{12}')X_{11}$$

and

$$\mathscr{I}_2 = X_{21}'(I_{n_2} - X_{22}(X_{22}'X_{22})^{-}X_{22}')X_{21} .$$

The term information matrix stems from the fact that the variance of the BLUE of an estimable function $p'\theta_1$ is proportional to $p'\mathscr{I}_1^{-}p$ and $p'\mathscr{I}_2^{-}p$ for the two experiments. Similarly the linear model and the information matrix associated with the $n_1 + n_2$ observations obtained from combining the two experiments may be written as

$$Y = \binom{Y_1}{Y_2} = \binom{X_{11}}{X_{21}}\theta_1 + \binom{X_{12}}{X_{22}}\theta_2 + \epsilon \ , \ E(\epsilon) = 0 \ , \ V(\epsilon) = \sigma^2 I$$

and

$$\mathscr{I} = (X_{11}' : X_{21}')[I_n - \binom{X_{12}}{X_{22}}(X_{12}'X_{12} + X_{22}'X_{22})^{-}(X_{12}' : X_{22}')]\binom{X_{11}}{X_{21}}$$

where $n = n_1 + n_2$.

For simplicity of notations, let us define the following orthogonal projection operators:

$$P_1 = P_{X_{12}} \ , \ P_2 = P_{X_{22}} \ , \ P = P_{\binom{X_{12}}{X_{22}}} \ ,$$

where the notation P_A was defined in Section 2. The information matrices may now be written as

$$\mathcal{I}_1 = X_{11}'(I_{n_1}-P_1)X_{11}$$
$$\mathcal{I}_2 = X_{21}'(I_{n_2}-P_2)X_{21}$$
$$\mathcal{I} = (X_{11}' : X_{21}')(I_n-P)\binom{X_{11}}{X_{21}}$$

We shall drop the subscripts for the identity matrices whenever they are understood from the context.

It is easy to verify that $\mathcal{I} \geq \mathcal{I}_1$ and $\mathcal{I} \geq \mathcal{I}_2$. Hence information is nondecreasing in the number of observations. In the following theorem we establish much more - information matrices are shown to be superadditive.

<u>Theorem 3.1.</u>
$$\mathcal{I} \geq \mathcal{I}_1 + \mathcal{I}_2 \tag{3.1}$$

with equality iff

$$\mathcal{R}\begin{pmatrix} X_{12}' & X_{11} \\ X_{22}' & X_{21} \end{pmatrix} \subset \mathcal{R}\begin{pmatrix} X_{12}' & X_{12} \\ X_{22}' & X_{22} \end{pmatrix} \tag{3.2}$$

<u>Proof</u>: $\mathcal{I} - \mathcal{I}_1 - \mathcal{I}_2 = (X_{11}' : X_{21}')(\mathrm{diag}(P_1, P_2) - P)\binom{X_{11}}{X_{21}}$

P is the orthogonal projection operator onto $\mathcal{R}\binom{X_{12}}{X_{22}}$, while

$\operatorname{diag}(P_1, P_2)$ is the orthogonal projection operator onto $\mathcal{M}\begin{pmatrix} X_{12} & 0 \\ 0 & X_{22} \end{pmatrix}$. Clearly,

$$\mathcal{M}\begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix} \subset \mathcal{M}\begin{pmatrix} X_{12} & 0 \\ 0 & X_{22} \end{pmatrix}.$$

Thus

$$(\operatorname{diag}(P_1, P_2))P = P$$

So, by Theorem 5.1.3 of Rao and Mitra (1971) $\operatorname{diag}(P_1, P_2) - P$ is the orthogonal projection operator onto

$$\left\{\mathcal{M}\begin{pmatrix} X_{12} & 0 \\ 0 & X_{22} \end{pmatrix}\right\} \cap \left\{\mathcal{M}\begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix}\right\}^{\perp} = \mathcal{S}(\text{say})$$

Here for any subspace L, $L^{\perp}$ will denote the space of vectors orthogonal to each vector in L.

In particular, $\operatorname{diag}(P_1, P_2) - P$ is nonnegative definite and hence $\mathcal{I} - \mathcal{I}_1 - \mathcal{I}_2 \geq 0$.

Equality is achieved iff

$$(\operatorname{diag}(P_1, P_2) - P)\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} = 0,$$

which is equivalent to

$$\mathcal{M}\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} \subset \mathcal{S}^{\perp} \qquad (3.3)$$

Note that $\mathcal{S}$ consists of vectors $(z_1' : z_2')'$ such that $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathcal{M}\begin{pmatrix} X_{12} & 0 \\ 0 & X_{22} \end{pmatrix}$ and $(X_{12}' : X_{22}')\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = 0$. i.e., $z_1 = X_{12}u_1$, $z_2 = X_{22}u_2$, and $X_{12}'X_{12}u_1 + X_{22}'X_{22}u_2 = 0$ for some u_1 and u_2. (3.3) is satisfied iff $X_{11}'z_1 + X_{21}'z_2 = 0$ for all z_1, z_2 of the above form. This is equivalent to the statement

$$X_{11}'X_{12}u_1 + X_{21}'X_{22}u_2 = 0 \text{ whenever } X_{12}'X_{12}u_1 + X_{22}'X_{22}u_2 = 0 \qquad (3.4)$$

(3.4) is clearly equivalent to (3.2). Hence the theorem.

Note that (3.2) and (3.4) are equivalent to:

$$X_{11}'X_{12}a_1 = X_{21}'X_{22}a_2 \text{ whenever } X_{12}'X_{12}a_1 = X_{22}'X_{22}a_2 \qquad (3.5)$$

These equations give geometric interpretation of equality in equation (3.1). In the next theorem we give a statistical interpretation. We shall show that every BLUE in the combined experiment is a sum of some suitably chosen BLUEs of the individual experiments. Each of these BLUEs may, however, be estimating different parametric functions.

Let $\mathcal{E}_i$ = {All BLUEs for all estimable functions of θ_1 in experiment i}, i = 1,2, and $\mathcal{E}$ = {All BLUEs for all estimable functions of θ_1 in the combined experiment}. By Lemma 2.2 (ii), $\mathcal{E}_i = \{b'Y \mid b \in \mathcal{M}[(I-P_i)X_{i1}]\}$, i = 1,2 and $\mathcal{E} = \{b'Y \mid b \in \mathcal{M}[(I-P)(X_{11}' : X_{21}')']\}$.

Theorem 3.2. $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ iff any $b'Y \in \mathcal{E}$ can be written as $b'Y = b_1'Y_1 + b_2'Y_2$ with $b_i'Y_i \in \mathcal{E}_i$, $i = 1,2$.

Proof: If $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$, then from the proof of Theorem 3.1 it follows that

$$(X_{11}' : X_{21}')(I-P) = (X_{11}' : X_{21}')(\mathrm{diag}(I-P_1,\ I-P_2)).$$

Hence $h'(X_{11}' : X_{21}')(I-P)Y = h'X_{11}'(I-P_1)Y_1 + h'X_{21}'(I-P_2)Y_2$. This establishes the necessity. To prove the sufficiency, suppose $b'Y \in \mathcal{E}$ and

$$b'Y = b_1'Y_1 + b_2'Y_2 = (b_1' : b_2')Y$$

where $b_1'Y_1 \in \mathcal{E}_1$ and $b_2'Y_2 \in \mathcal{E}_2$. Applying Lemma 2.2 (i) to the combined model,

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathcal{M}\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

i.e., for some h, h_1

$$b_1 = X_{11}h + X_{12}h_1 \tag{3.6}$$

$$b_2 = X_{21}h + X_{22}h_1 \tag{3.7}$$

Premultiplying (3.6) by $(I-P_1)$ we get

$$(I-P_1)b_1 = (I-P_1)X_{11}h$$

But $(I-P_1)b_1 = b_1$, since $b_1 \in \mathcal{M}[(I-P_1)X_{11}]$ and $I-P_1$ is idempotent. Hence

$$b_1 = (I-P_1)X_{11}h\,.$$

Similarly $b_2 = (I-P_2)X_{21}h$, from equation (3.7).
So $b'Y = h'X_{11}'(I-P_1)Y_1 + h'X_{21}'(I-P_2)Y_2$. Since $b'Y$ can be any element in $\mathcal{E}$, this implies that for each s, there exists an h (a function of s) such that

$$s'(X_{11}' : X_{21}')(I-P)Y = h'X_{11}'(I-P_1)Y_1 + h'X_{21}'(I-P_2)Y_2 \qquad (3.8)$$

Taking expectation on both sides of (3.8), we obtain

$$s'\mathcal{J} = h'\mathcal{J}_1 + h'\mathcal{J}_2\,.$$

Defining
$$\mathcal{J}_o = \mathcal{J}_1 + \mathcal{J}_2 \qquad (3.9)$$

$$\mathcal{J}s = \mathcal{J}_o h\,. \qquad (3.10)$$

Since (3.10) is true for all s, $\mathcal{J} = \mathcal{J}_o H$, for some matrix H.

Thus $\mathcal{M}(\mathcal{J}) \subset \mathcal{M}(\mathcal{J}_o)$. But $\mathcal{M}(\mathcal{J}_o) \subset \mathcal{M}(\mathcal{J})$ from (3.1). Hence we have,

$$\mathcal{M}(\mathcal{J}) = \mathcal{M}(\mathcal{J}_o) \qquad (3.11)$$

From (3.10) we get,

$$h = \mathcal{J}_o^- \mathcal{J} s \quad \text{for some } \mathcal{J}_o^- .$$

Now taking variances in (3.8), it follows that

$$\begin{aligned} s'\mathcal{J}s &= h'\mathcal{J}_1 h + h'\mathcal{J}_2 h = h'\mathcal{J}_o h = s'\mathcal{J}\mathcal{J}_o^{-'}\mathcal{J}_o \mathcal{J}_o^- \mathcal{J}s \\ &= s'\mathcal{J}\mathcal{J}_o^-\mathcal{J}s , \end{aligned}$$

since $\mathcal{J}\mathcal{J}_o^{-'}\mathcal{J}_o = \mathcal{J}\mathcal{J}_o^-\mathcal{J}_o = \mathcal{J}$, using the fact that $\mathcal{J}_o^{-'}$ is a g-inverse of $\mathcal{J}_o$, the latter being symmetric, and using (3.11). The above being true for all s, and since both $\mathcal{J}$ and $\mathcal{J}\mathcal{J}_o^-\mathcal{J}$ are symmetric, one may write

$$\mathcal{J}\mathcal{J}_o^-\mathcal{J} = \mathcal{J} \qquad (3.12)$$

Note that this relation is true for all g-inverses of $\mathcal{J}_o$ since $\mathcal{J}\mathcal{J}_o^-\mathcal{J}$ is invariant under choice of $\mathcal{J}_o^-$ (see Rao and Mitra (1971) Lemma 2.2.4(iii)).

Let T be a nonsingular matrix such that $\mathscr{I} = T' \operatorname{diag}(D,0)T$ and $\mathscr{I}_o = T'\operatorname{diag}(D_o,0)T$. D and D_o have the same dimension and are both nonsingular due to (3.11). Choose $\mathscr{I}_o^- = T^{-1}\operatorname{diag}(D_o^{-1},0)T'^{-1}$. Using these representations in (3.12) we obtain $D = D_o$ and hence $\mathscr{I} = \mathscr{I}_o$.

In view of Theorem 3.2 , the results of Theorem 3.1 can be given the following interpretation. $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$, iff

$$\mathscr{E} = \mathscr{E}_1 \oplus \mathscr{E}_2 \tag{3.13}$$

the algebraic sum of $\mathscr{E}_1$ and $\mathscr{E}_2$. If $\mathscr{I} > \mathscr{I}_1 + \mathscr{I}_2$, then (3.13) is violated. This means that elements in $\mathscr{E} - \mathscr{E}_1 \oplus \mathscr{E}_2$ are BLUEs which cannot be computed from BLUEs of the original experiment. This is surely due to one or both of the following conditions:

<u>C.1</u>. (Expectation condition): There exists a $p \in \mathscr{L}(\mathscr{I})$ which cannot be written as $p'\theta_1 = p_1'\theta_1 + p_2'\theta_1$ with $p_1 \in \mathscr{L}(\mathscr{I}_1)$ and $p_2 \in \mathscr{L}(\mathscr{I}_2)$. Note that the column space of an information matrix provides all estimable functions.

<u>C.2</u>. (Variance condition): There exist a p in the set $\{p \mid p \in \mathscr{L}(\mathscr{I})$ and $p'\theta_1 = p_1'\theta_1 + p_2'\theta_1$, where $p_1 \in \mathscr{L}(\theta_1)$ and $p_2 \in \mathscr{L}(\mathscr{I}_2)$ for some $p_1 , p_2\}$, but for all representations of such a p, the BLUE of $p'\theta_1$ in combined experiment has a smaller

variance than the sum of the BLUE of $p_1'\theta_1$ in experiment 1 and the BLUE of $p_2'\theta_1$ in experiment 2.

We shall elucidate the above with some examples in block designs. First we state a corollary which gives a general characterization of $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ for block designs.

We assume the usual additive model. θ_1 consists of the treatment effects and θ_2 the block effects. Let experiment 1 use blocks $B_1,\ldots,B_q,B_{q+1},\ldots,B_r$ and let experiment 2 use blocks $B_{q+1},\ldots,B_r,B_{r+1},\ldots,B_{r+s}$. To keep our discussions general, we assume that the number of observations in B_i for experiment 1 is k_{1i} and that for experiment 2 is k_{2i}, and k_{1i} need not equal k_{2i} when $q + 1 \le i \le r$. Let N_i, $i = 1,2$ denote the incidence matrices for the two experiments, i.e.,

$$N_1 = X_{11}'X_{12} \quad \text{and} \quad N_2 = X_{21}'X_{22} .$$

X_{12} and X_{22} has $r+s$ columns each. Moreover partition

$$N_1 = [N_{11} : N_{12} : 0] \quad \text{and} \quad N_2 = [0 : N_{22} : N_{23}]$$

where N_{11} has q columns, N_{12} and N_{22} have $r-q$ columns each, N_{23} has s columns. It is clear that $X_{12}'X_{12} = \text{diag}(k_{11},\ldots,k_{1r},0,\ldots,0)$, $X_{22}'X_{22} = \text{diag}(0,\ldots,0,k_{2q+1},\ldots,k_{2r+s})$. The corollary may now be stated.

<u>Corollary 3.1</u>. For a block design $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$ iff

$$(\mathrm{diag}(k_{1q+1}^{-1},\ldots,k_{1r}^{-1}))N_{12}' = (\mathrm{diag}(k_{2q+1}^{-1},\ldots,k_{2q+1}^{-1}))N_{22}' \quad (3.13)$$

<u>Proof</u>: From equation (3.2), $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$ iff there exists a matrix T such that

$$N_1' = X_{12}'X_{12}T \quad \text{and} \quad N_2' = X_{22}'X_{22}T$$

Partitioning $T' = [T_1' : T_2' : T_3']$, one obtains

$$N_{11}' = (\mathrm{diag}(k_{11},\ldots,k_{1q}))T_1 \quad , \quad N_{12}' = (\mathrm{diag}(k_{1q+1},\ldots,k_{1r}))T_2$$
$$N_{22}' = (\mathrm{diag}(k_{2q+1},\ldots,k_{2r}))T_2 \; , \; N_{23}' = (\mathrm{diag}(k_{2r+1},\ldots,k_{2r+s}))T_3 \; .$$

Clearly T_1 and T_3 always exists, and T_2 exists iff (3.13) is satisfied.

Thus $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$ iff the designs in the common blocks are "proportional" in the sense of (3.13). In particular, if $k_{1i} = k_{2i}$, $q + 1 \le i \le r$, then the designs should be isomorphic, i.e., $N_{12}' = N_{22}'$. If the two experiments have no blocks in common then obviously $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$ -- a fact which is generally well known to researchers in the theory of optimal designs.

<u>Example 3.1</u>. This is a simple though somewhat pathological example. However, one may easily extend it to more nontrivial

cases. Consider the following block designs with incidence matrices

$$N_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

the two experiments using the same treatments and blocks. The combined experiment has the incidence matrix

$$N = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} .$$

Here $R(\mathscr{I}) = r(\mathscr{I}_1) = r(\mathscr{I}_2)$ --there is only one estimable function. Moreover $\mathscr{I}_1 = \mathscr{I}_2$. But $\mathscr{I} > \mathscr{I}_1 + \mathscr{I}_2$, by Corollary 3.1. This is because each experiment has an observation confounded with the second block. These observations are released for estimation of treatments when the experiments are combined. This is an example of condition C2.

Example 3.2. Consider the following experiments using the same treatments and blocks.

$$N_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Naming the treatment effects τ_i, $1 \le i \le 4$, we see that the estimable functions in each experiment are $\tau_1 - \tau_2$ and $\tau_3 - \tau_4$ and their linear combinations only. $\tau_2 - \tau_3$ becomes estimable

too upon combining the experiments. Here $\mathscr{I}_1 = \mathscr{I}_2$ but $\mathscr{I} > \mathscr{I}_1 + \mathscr{I}_2$. It will be seen in the next section that the BLUEs for $\tau_1 - \tau_2$ and $\tau_3 - \tau_4$ for the combined experiment is a combination of the BLUEs of the individual experiments. Thus this is an example of condition C1. A combination of designs in Example 1 and Example 2 will give an example of both C1 and C2.

In view of Corollary 3.1 it is easy to construct examples where $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$. Note also that $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$ when there are no nuisance parameters, i.e., $\theta_2 = 0$.

4. ON THE ADDITIVITY OF BLUES IN COMBINED EXPERIMENTS.

If two sets of data are collected for a fixed linear model then it is obvious that one should combine the data and carry out one analysis. In Section 3 we measured the gain in combining the data in terms of information matrices. In this section we shall study the status of the Gauss-Markov theorem for the combined data and explore its relation to the associated Gauss-Markov theorems for the individual sets of data. In particular, we shall deal explicitly with the following questions:

Question. 1. Suppose we are given a linear parametric function $p'\theta_1$ which is estimable under both sets of data. Under what conditions,

BLUE of $p'\theta_1$ under the combined data = linear combination of BLUEs of $p'\theta_1$ under the individual sets of data? (4.1)

Question. 2. Same as Question 1, except that we insist that the statement (4.1) is true for all estimable functions. Thus if ϕ is any linear function of θ_1 estimable under the both sets of data, we want

BLUE of ϕ under the combined data = linear combination of BLUEs of ϕ under individual sets of data. (4.2)

We shall completely characterize cases under which (4.1) and (4.2) are valid. These characterizations turn out to be mathematically interesting and reveal more about the structure of the information matrices. These results would be of practical significance since in the cases where (4.1) or (4.2) are valid, a lot of computational time can be saved if the BLUEs for each individual sets of data are already available.

Before starting on these problems, we remark that Theorem 3.2 answers a question similar to Question 2. If $\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2$, h is a solution of

$$\mathscr{I} h = p , \tag{4.3}$$

and $p_i = \mathscr{I}_i h$, $i = 1,2$ then from the proof of the necessity of Theorem 3.2, (BLUE $p'\theta_1$ in combined data) = (BLUE $p_1'\theta_1$ in experiment 1) + (BLUE $p_2'\theta_1$ in experiment 2). Note that p_1 and p_2 are invariant whatever h, satisfying (4.3). However, $p_1'\theta_1$ and $p_2'\theta_1$ may not be multiples of $p'\theta_1$.

Let us look at Question 1. Suppose $b_i'Y_i$ is the BLUE of $p'\theta_1$ in experiment i, $i = 1,2$. Here $p \in \mathcal{L}(\mathcal{I}_1) \cap \mathcal{L}(\mathcal{I}_2)$. If $\alpha_1 b_1'Y_1 + \alpha_2 b_2'Y_2$ is the BLUE of $p'\theta_1$ when the data are combined, the unbiasedness condition, $p'\theta_1 = E(\alpha_1 b_1'Y_1 + \alpha_2 b_2'Y_2)$, for all θ_1, implies that $\alpha_1 + \alpha_2 = 1$. Among all such unbiased estimators the one with minimum variance has

$$\alpha_1 = \frac{V(b_2'Y_2)}{V(b_1'Y_1) + V(b_2'Y_2)} = \alpha(\text{say}) \quad \text{and} \quad \alpha_2 = 1 - \alpha .$$

Explicitly,

$$\alpha = \frac{p'\mathcal{I}_2^- p}{p'\mathcal{I}_1^- p + p'\mathcal{I}_2^- p} \tag{4.4}$$

So the only linear combination which can be BLUE of $p'\theta_1$ under the combined experiment is

$$b'Y = \alpha b_1'Y_1 + (1-\alpha)b_2'Y_2 . \tag{4.5}$$

The following Theorem answers Question 1.

<u>Theorem 4.1</u>. Let $p \in \mathcal{L}(\mathcal{I}_1) \cap \mathcal{L}(\mathcal{I}_2)$ and $p \neq 0$. Then equation (4.1) is satisfied iff there exists vectors h_1 and h_2 such that

$$\mathcal{I}_1 h_1 = \alpha p \quad , \quad \mathcal{I}_2 h_1 = (1-\alpha)p \tag{4.6}$$

$$X_{12}'X_{11}h_1 + X_{12}'X_{12}h_2 = 0 \;\; , \;\; X_{22}'X_{21}h_1 + X_{22}'X_{22}h_2 = 0 \qquad (4.7)$$

Proof: $b'Y$ in (4.5) is obviously unbiased. By Lemma 2.2 (i) it is BLUE of $p'\theta_1$ iff

$$\begin{pmatrix} \alpha b_1 \\ (1-\alpha)b_2 \end{pmatrix} \in \mathcal{M}\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

i.e., iff there exists h_1 and h_2 such that

$$\alpha b_1 = X_{11}h_1 + X_{12}h_2 \qquad (4.8)$$

$$(1-\alpha)b_2 = X_{21}h_1 + X_{22}h_2 \qquad (4.9)$$

(4.8) can be equivalently written as

$$\alpha b_1 = (I-P_1)X_{11}h_1 + P_1X_{11}h_1 + X_{12}h_2 \qquad (4.10)$$

where $P_1 = X_{12}(X_{12}'X_{12})^-X_{12}'$, as in Section 3. Premultiplying (4.10) by X_{12}' we get, since $X_{12}'b_1 = 0$,

$$0 = X_{12}'X_{11}h_1 + X_{12}'X_{12}h_2 \,. \qquad (4.11)$$

Premultiplying (4.11) by $X_{12}(X_{12}'X_{12})^-$,

$$0 = P_1X_{11}h_1 + X_{12}h_2 \,, \qquad (4.12)$$

This, together with (4.10) implies

$$\alpha b_1 = (I-P_1)X_{11}h_1 \tag{4.13}$$

It is easy to see the equivalence of (4.11) and (4.12). Therefore (4.11) and (4.13) together imply (4.10). Hence (4.10) is equivalent to $\{(4.11), (4.13)\}$. Premultiplying (4.13) by X_{11}', we have, since $X_{11}'b_1 = p$,

$$\alpha p = \mathscr{I}_1 h_1 \tag{4.14}$$

By Lemma 2.2 (ii)

$$b_1 = (I-P_1)X_{11}t$$

for some t. Thus $p = X_{11}'(I-P_1)X_{11}t$. So (4.14) becomes

$$\alpha X_{11}'(I-P_1)X_{11}t = X_{11}'(I-P_1)X_{11}h_1 ,$$

which implies that

$$\alpha (I-P_1)X_{11}t = (I-P_1)X_{11}h_1 .$$

Thus (4.13) and (4.14) are equivalent. Hence (4.8) is equivalent to

$$X_{12}'X_{11}h_1 + X_{12}'X_{12}h_2 = 0$$

and

$$\mathcal{J}_1 h_1 = \alpha p$$

Similarly (4.9) is equivalent to

$$\mathcal{J}_2 h_1 = (1-\alpha)p \quad \text{and} \quad X_{22}'X_{21}h_1 + X_{22}'X_{22}h_2 = 0 \; .$$

Hence the theorem.

Let us move on to Question 2. Here, we want (4.1) to be valid for all p in $\mathcal{R}(\mathcal{J}_1) \cap \mathcal{R}(\mathcal{J}_2)$ which we assume to be different from $\{0\}$. Let T be a nonsingular matrix such that,

$$\mathcal{J}_1 = T'\mathrm{diag}(D_{11}, D_{12}, 0, 0)T \; , \; \mathcal{J}_2 = T'\mathrm{diag}(D_{21}, 0, D_{22}, 0)T \qquad (4.15)$$

where the D_{ij}'s are diagonal matrices. Note that the simultaneous diagonalization of $\mathcal{J}_1$ and $\mathcal{J}_2$ are carried out in such a way that the corresponding component matrices in $\mathrm{diag}(D_{11}, D_{12}, 0, 0)$ and $\mathrm{diag}(D_{21}, 0, D_{22}, 0)$ are of the same dimension. D_{11} and D_{21} are nonsingular. If $\mathcal{R}(\mathcal{J}_1) \subset \mathcal{R}(\mathcal{J}_2)$, then there will be no D_{12} and we shall write the matrices in (4.15) with only three diagonal blocks. Similar modifications are needed when $\mathcal{R}(\mathcal{J}_2) \subset \mathcal{R}(\mathcal{J}_1)$. If $\mathcal{R}(\mathcal{J}_1) = \mathcal{R}(\mathcal{J}_2)$, then we need only two diagonal blocks, i.e.,

$$\mathcal{J}_1 = T'\mathrm{diag}(D_{11}, 0)T \; , \; \mathcal{J}_2 = T'\mathrm{diag}(D_{21}, 0)T \qquad (4.16)$$

Considering only the general case when $\mathcal{R}(\mathcal{J}_1) \cap \mathcal{R}(\mathcal{J}_2)$ contains neither $\mathcal{R}(\mathcal{J}_1)$ nor $\mathcal{R}(\mathcal{J}_2)$, we may insist that D_{12} and D_{22}

are nonsingular. The results for the special cases mentioned above may be obtained from the general results with obvious modifications - we shall only consider the case $\mathcal{M}(\mathcal{I}_1) = \mathcal{M}(\mathcal{I}_2)$ separately, as this is of special interest.

Partitioning, $T' = (T_1' : T_2' : T_3' : T_4')$, we get from (4.15),

$$\mathcal{I}_1 = T_1'D_{11}T_1 + T_2'D_{12}T_2 \ , \ \mathcal{I}_2 = T_1'D_{21}T_1 + T_3'D_{22}T_3$$

Defining

$$\mathcal{I}_{11} = T_1'D_{11}T_1 \ , \mathcal{I}_{12} = T_2'D_{12}T_2 \ , \mathcal{I}_{21} = T_1'D_{21}T_1 \ , \mathcal{I}_{22} = T_3'D_{22}T_3$$

we obtain

$$\mathcal{I}_1 = \mathcal{I}_{11} + \mathcal{I}_{12} \ , \ \mathcal{I}_2 = \mathcal{I}_{21} + \mathcal{I}_{22} \tag{4.17}$$

$$\begin{gathered} \mathcal{M}(\mathcal{I}_{11}) \cap \mathcal{M}(\mathcal{I}_{12}) = \{0\} \ , \ \mathcal{M}(\mathcal{I}_{21}) \cap \mathcal{M}(\mathcal{I}_{22}) = \{0\} \\ \mathcal{M}(\mathcal{I}_{12}) \cap \mathcal{M}(\mathcal{I}_{22}) = \{0\} \ , \ \mathcal{M}(\mathcal{I}_{11}) = \mathcal{M}(\mathcal{I}_{21}) = \mathcal{M}(\mathcal{I}_1) \cap \mathcal{M}(\mathcal{I}_2) \ . \end{gathered} \tag{4.18}$$

From Theorem 3.1 it follows that

$$\mathcal{I} = \mathcal{I}_{11} + \mathcal{I}_{12} + \mathcal{I}_{21} + \mathcal{I}_{22} + R$$

where R is a nonnegative definite matrix.

<u>Theorem 4.2</u>. BLUE of $p'\theta_1$ in the combined experiment

$$= \alpha(\text{BLUE of } p'\theta_1 \text{ in experiment 1})$$

$$+ (1-\alpha)(\text{BLUE of } p'\theta_1 \text{ in experiment 2}) \qquad (4.19)$$

for all $p \in \mathcal{R}(\mathcal{J}_1) \cap \mathcal{R}(\mathcal{J}_2)$, iff

$$\mathcal{J}_{21} = \lambda \mathcal{J}_{11} \text{ for some } \lambda > 0 \qquad (4.20)$$

and

$$\mathcal{J} = \mathcal{J}_{11} + \mathcal{J}_{21} + R_o \qquad (4.21)$$

with

$$R_o = \mathcal{J}_{12} + \mathcal{J}_{22} + R \text{ satisfying}$$

$$\mathcal{R}(\mathcal{J}_{11}) \cap \mathcal{R}(R_o) = \{0\} \qquad (4.22)$$

<u>Proof</u>: Suppose (4.19) is given. Then (4.6) is satisfied for each $p \in \mathcal{R}(\mathcal{J}_{11}) = \mathcal{R}(\mathcal{J}_{21})$. The equation

$$\mathcal{J}_1 h_1 = \alpha p$$

can be written equivalently as

$$\mathcal{J}_{11} h_1 = \alpha p \quad , \quad \mathcal{J}_{12} h_1 = 0$$

using the facts that $p \in \mathcal{R}(\mathcal{J}_{11})$ and $\mathcal{R}(\mathcal{J}_{11}) \cap \mathcal{R}(\mathcal{J}_{12}) = \{0\}$.

So, equation (4.6) implies

$$\mathcal{S}_{11} h_1 = \alpha p \quad \text{and} \quad \mathcal{S}_{21} h_1 = (1-\alpha) p \ ,$$

and this is true for all $p \in \mathcal{M}(\mathcal{S}_{11})$. Lemma 2.3 now gives

$$\mathcal{S}_{21} = \lambda \mathcal{S}_{11} \quad , \text{ for some } \lambda > 0$$

which is (4.20). Observe that,

$$\alpha = \frac{p'\mathcal{S}_2^- p}{p'\mathcal{S}_1^- p + p'\mathcal{S}_2^- p} = \frac{p'\mathcal{S}_{21}^- p}{p'\mathcal{S}_{11}^- p + p'\mathcal{S}_{21}^- p} = \frac{1}{1+\lambda} \tag{4.23}$$

since by Lemma 2.4, $p'\mathcal{S}_i^- p = p'\mathcal{S}_{i1}^- p$, $i = 1,2$. Thus α is independent of p.

Taking variances on both sides of equation (4.19), one obtains, for each $p \in \mathcal{M}(\mathcal{S}_{11})$,

$$\begin{aligned} p'\mathcal{S}^- p &= \alpha^2 p'\mathcal{S}_1^- p + (1-\alpha)^2 p'\mathcal{S}_2^- p \\ &= \alpha^2 p'\mathcal{S}_{11}^- p + (1-\alpha)^2 p'\mathcal{S}_{21}^- p \\ &= (1+\lambda)^{-1} p'\mathcal{S}_{11}^- p \ , \quad \text{using (4.23)} \\ &= p'\mathcal{S}_o^- p \ , \end{aligned}$$

where $\mathcal{S}_o = \mathcal{S}_{11} + \mathcal{S}_{21} = (1+\lambda)\mathcal{S}_{11}$. Since this is true for all $p \in \mathcal{M}(\mathcal{S}_o)$, we get

$$\sigma^{-2}\, V(\text{BLUE of } p'\theta_1 \text{ in combined experiment})$$
$$= p'\mathcal{I}^- p$$
$$= p'\mathcal{I}_o^- p$$

using the "(i) $\Rightarrow$ (ii)" part of Lemma 2.4, and (4.22). This being the same as (4.24), we appeal to the uniqueness of the BLUE in the classical Gauss-Markov model to conclude that the estimator $b'Y$ is the BLUE in the combined model. Hence we get equation (4.19).

<u>Remark 4.1</u>. Under conditions of the Theorem,

$$\mathcal{I} = \mathcal{I}_o + R_o \;.$$

Using Lemma 2.4 and equation (4.22), we get $p'\mathcal{I}^- p = p'\mathcal{I}_o^- p$, for each $p \in \mathcal{M}(\mathcal{I}_o)$ and $p'\mathcal{I}^- p = p'R_o^- p$, for each $p \in \mathcal{M}(R_o)$. In this sense $\mathcal{I}_o$ and R_o are themselves information matrices. Thus (4.21) gives a decomposition of $\mathcal{I}$ into two information matrices whose spaces are disjoint. In a sense the combined experiment can be looked upon as the union of two disjoint (fictitious) experiments, one consisting of the common part of the individual experiments (corresponds to $\mathcal{I}_o$), and the other consisting of the remainder of the individual experiments (corresponds to $\mathcal{I}_{12}$ and $\mathcal{I}_{22}$) and the portion which is purely the profit from combining (corresponding to R). Also note that (4.22) implies

$$\mathcal{M}(\mathcal{I}_{11}) \cap \mathcal{M}(R) = \{0\} .$$

<u>Remark 4.2</u>. Since $\mathcal{I}_1 = \mathcal{I}_{11} + \mathcal{I}_{12}$ and $\mathcal{M}(\mathcal{I}_{11}) \cap \mathcal{M}(\mathcal{I}_{12}) = \{0\}$, $\mathcal{I}_{11}$ and $\mathcal{I}_{21}$ are themselves information matrices, by Lemma 2.4. $\mathcal{I}_{11}$ is the information matrix of parametric functions common to both experiments and $\mathcal{I}_{12}$ that of function which can be estimated from experiment 1 only. A similar explanation holds for $\mathcal{I}_{21}$ and $\mathcal{I}_{22}$. Thus when the conditions of the Theorem are valid, estimable functions in $\mathcal{M}(\mathcal{I}_1) \cap \mathcal{M}(\mathcal{I}_2)$ are estimated by a linear combination of the BLUEs. Estimable function from $\mathcal{M}(\mathcal{I}_{12})$ are nonestimable in experiment 2, but thir BLUEs may use observations from experiment 2 when the experiments are combined. This is because $\mathcal{M}(\mathcal{I}_{12}) \cap \mathcal{M}(R)$ need not be $\{0\}$. A similar observation holds for functions from $\mathcal{M}(\mathcal{I}_{22})$.

As an example, consider the following block designs, given by their incidence matrices

$$N_1' = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad N_2' = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

The combined experiment has the incidence matrix

$$N' = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix}$$

Straightforward computations show that $\mathcal{I}_{11} = \mathcal{I}_{21} = \mathrm{diag}(0,0,A)$, where $A = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, and

$$\mathcal{J} = \frac{1}{4}\begin{bmatrix} 5 & -3 & -1 & -1 & 0 & 0 \\ -3 & 5 & -1 & -1 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ -1 & -1 & -1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & 0 & -4 & 4 \end{bmatrix}$$

Clearly the conditions of the theorem are satisfied. Hence the BLUE of $\tau_5 - \tau_6$ is a linear combination of the BLUEs in the original experiments. $\tau_3 - \tau_4$ can be estimated from experiment 2 only. $\tau_1 - \tau_2$ is estimable in experiment 1 only, but its BLUE in the combined experiment uses observations from experiment 2 also. This is due to the fact that the observations in block 2 in each individual experiment cannot be used due to confounding, but are released when they are combined. $\tau_2 - \tau_3$ can be estimated only in the combined experiment. We give this somewhat oversimplefied example to illustrate Theorem 4.2 in a way that can be seen obviously without getting involved into complicated structures of the matrices associated with problem.

In the following theorem, the conditions of Theorem 4.2 are expressed in a form which is better suited for computational verification. We give a general form which must be simplified in particular settings in which an experimenter is interested.

<u>Theorem 4.3</u>. Equation (4.19) is satisfied for all p in $\mathcal{L}(\mathcal{J}_1) \cap \mathcal{L}(\mathcal{J}_2)$ iff

$$\mathcal{J}_{21} = \lambda \mathcal{J}_{11}, \text{ for some } \lambda > 0. \tag{4.20}$$

and

$$\mathcal{L}\begin{pmatrix} X_{12}'X_{11}Q_1' \\ X_{22}'X_{21}Q_1' \end{pmatrix} \subset \mathcal{L}\begin{pmatrix} X_{12}'X_{12} & X_{12}'X_{11}Q_4' \\ X_{22}'X_{22} & X_{22}'X_{21}Q_4' \end{pmatrix} \tag{4.25}$$

where $T^{-1} = Q' = (Q_1' : Q_2' : Q_3' : Q_4')$, T being defined in (4.15), and the partitions corresponding to those of T'.

Proof: As noted in the proof of Theorem 4.2, (4.6) is equivalent to

$$\mathcal{J}_{11}h_1 = \alpha p \ , \ \mathcal{J}_{21}h_1 = (1-\alpha)p \qquad (4.26)$$

and

$$\mathcal{J}_{12}h_1 = 0 \ , \ \mathcal{J}_{22}h_1 = 0 \qquad (4.27)$$

since $p \in \mathcal{L}(\mathcal{J}_{11})$. Equation (4.26) can be replaced by

$$\mathcal{J}_{21} = \lambda \mathcal{J}_{11} , \text{ for some } \lambda > 0 , \ \mathcal{J}_{11}h_1 = \alpha p \qquad (4.28)$$

So (4.19) is equivalent to the following statement:

For each $p \in \mathcal{L}(\mathcal{J}_{11})$, there exists h_1 , h_2 such that (4.27), (4.28) and (4.7) are satisfied. (4.29)

Note that (4.28) must imply that $\alpha = (1+\lambda)^{-1}$. Let $Th_1 = s$, $(T')^{-1}p = q$ and partition $s' = (s_1' : s_2' : s_3' : s_4')$ and $q' = (q_1' : q_2' : q_3' : q_4')$. As in the proof of Lemma 2.3, q_2 , q_3 and q_4 are all null vectors since $p \in \mathcal{L}(\mathcal{J}_{11})$.

Ovserve that

$$\mathcal{J}_{11}h_1 = \alpha p \Leftrightarrow D_{11}s_1 = \alpha q_1 \Leftrightarrow s_1 = \alpha D_{11}^{-1}q_1$$

The first equivalence follows since $Q_1T_1' = I$, $T_j'Q_jT_j' = T_j'$ and hence $T_1'Q_1p = p$, since $p \in \mathcal{N}(T_1')$. Similarly

$$\mathcal{F}_{12}h_1 = 0 \iff D_{12}s_2 = 0 \iff s_2 = 0$$

$$\mathcal{F}_{22}h_1 = 0 \iff D_{22}s_3 = 0 \iff s_3 = 0$$

Also,
$$X_{12}'X_{11}h_1 + X_{12}'X_{12}h_2 = 0$$

$$\iff \quad -X_{12}'X_{12}h_2 = X_{12}'X_{11}T^{-1}s = X_{12}'X_{11}Q_1's_1 + X_{12}'X_{11}Q_4's_4$$

$$\iff \quad X_{12}'X_{11}Q_1's_1 = -X_{12}'X_{12}h_2 - X_{12}'X_{11}Q_4's_4 \tag{4.30}$$

And similarly $X_{22}'X_{21}h_1 + X_{22}'X_{22}h_2 = 0$

$$\iff X_{22}'X_{21}Q_1's_1 = -X_{22}'X_{22}h_2 - X_{22}'X_{21}Q_4's_4 \tag{4.31}$$

So, (4.29) is equivalent to

> For each $p \in \mathcal{N}(\mathcal{F}_{11})$, there exists h_2 and s_4 such that $\mathcal{F}_{21} = \lambda\mathcal{F}_{11}$, for some $\lambda > 0$, and (4.30) and (4.31) are satisfied. (4.32)

p varies over $\mathcal{N}(\mathcal{F}_{11})$ iff q_1 is any arbitrary vector, which is true iff $s_1 = (1+\lambda)^{-1}D_{11}^{-1}q_1$ is any arbitrary vector.

s_2 and s_3 must be 0 as we have already noted. So, (4.32) is equivalent to

$\mathscr{I}_{21} = \lambda \mathscr{I}_{11}$, for some $\lambda > 0$, and there exists matrices H and S such that

$$X_{12}'X_{11}Q_1' = X_{12}'X_{12}H + X_{12}'X_{11}Q_4'S$$
$$X_{22}'X_{21}Q_1' = X_{22}'X_{22}H + X_{22}'X_{21}Q_4'S \,. \tag{4.33}$$

Clearly (4.33) is equivalent to (4.20) and (4.25).

A case of special interest is when both the experiments are designed to have the same estimable functions of θ_1. i.e.,

$$\mathscr{M}(\mathscr{I}_1) = \mathscr{M}(\mathscr{I}_2) \tag{4.34}$$

In this case $\mathscr{I}_{12} = \mathscr{I}_{22} = 0$, $\mathscr{I}_1 = \mathscr{I}_{11}$, $\mathscr{I}_2 = \mathscr{I}_{21}$, $R_o = R$. Theorems 4.2 and 4.3 can be restated with some simplifications when (4.34) is satisfied. We do these in Corollaries 4.1 and 4.2.

<u>Corollary 4.1</u>. If $\mathscr{M}(\mathscr{I}_1) = \mathscr{M}(\mathscr{I}_2)$, then for each $p \in \mathscr{M}(\mathscr{I}_1)$, BLUE of $p'\theta_1$ in the combined experiment is a linear combination of the BLUEs of $p'\theta_1$ in the individual experiments iff

$$\mathscr{I}_2 = \lambda \mathscr{I}_1 \quad \text{for some} \quad \lambda > 0$$

and

$$\mathscr{I} = \mathscr{I}_1 + \mathscr{I}_2 + R \quad \text{with} \quad \mathscr{M}(\mathscr{I}_1) \cap \mathscr{M}(R) = \{0\} \,.$$

<u>Corollary 4.2</u>. If $\mathcal{M}(\mathcal{I}_1) = \mathcal{M}(\mathcal{I}_2)$ then equation (4.19) is satisfied for all $p \in \mathcal{M}(\mathcal{I}_1)$ iff

$$\mathcal{I}_2 = \lambda \mathcal{I}_1 \quad \text{for some} \quad \lambda > 0$$

and

$$\mathcal{M}\begin{pmatrix} X_{12}'X_{11}U_1' \\ X_{22}'X_{21}U_1' \end{pmatrix} \subset \mathcal{M}\begin{pmatrix} X_{12}'X_{12} & X_{12}'X_{11}U_2' \\ X_{22}'X_{22} & X_{22}'X_{21}U_2' \end{pmatrix}$$

where $U' = (U_1' : U_2')$ is the inverse of T, which is any non-singular matrix diagonalizing $\mathcal{I}_1$, i.e.,

$$\mathcal{I}_1 = T' \operatorname{diag}(D_1, 0)T ,$$

and $T' = (T_1' : T_2')$. Both T_1' and U_1' have as many columns as $r(\mathcal{I}_1)$.

As illustrations consider the examples in Section 3. In Example 3.1, $\mathcal{I}_1 = \mathcal{I}_2$ but the conditions of Corollary 4.2 are violated. Hence the BLUE of $\tau_1 - \tau_2$ is not a linear combination of BLUEs of the individual experiment. Clearly, the observations in Block 2 in the individual experiments can only be used when the data are combined.

In Example 3.2, $\mathcal{I}_1 = \mathcal{I}_2$ and the conditions of Corollary 4.2 are satisfied. The BLUEs of $\tau_1 - \tau_2$ and $\tau_3 - \tau_4$ are obtained as linear combinations from the individual experiments. However $\tau_2 - \tau_3$ is estimable only when the data are combined.

Thus to obtain its BLUE the entire data has to be put together and analysed. This leads us to the next question. Under what conditions is (4.1) satisfied for all $p \in \mathcal{M}(\mathcal{I})$, if $\mathcal{M}(\mathcal{I}) = \mathcal{M}(\mathcal{I}_1) = \mathcal{M}(\mathcal{I}_2)$. The following corollary answers this as a special case of Corollary 4.1.

Corollary 4.3. If $\mathcal{M}(\mathcal{I}_1) = \mathcal{M}(\mathcal{I}_2) = \mathcal{M}(\mathcal{I})$, then for each $p \in \mathcal{M}(\mathcal{I})$, the BLUE of $p'\theta_1$ in the combined experiment is a linear combination of the BLUEs of $p'\theta_1$ in the individual experiments iff

$$\mathcal{I}_2 = \lambda \mathcal{I}_1 \quad \textit{for some} \quad \lambda > 0$$

and

$$\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2 .$$

Proof: Follows from Corollary 4.1 since if $\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2 + R$, and $\mathcal{M}(\mathcal{I}) = \mathcal{M}(\mathcal{I}_1) = \mathcal{M}(\mathcal{I}_2)$, then $\mathcal{I}_2 = \lambda \mathcal{I}_1$ and $\mathcal{M}(\mathcal{I}_1) \cap \mathcal{M}(R) = \{0\}$ iff $\mathcal{I}_2 = \lambda \mathcal{I}_1$ and $R = 0$.

For the sake of completeness, we state the version of Corollary 4.2 when $\mathcal{M}(\mathcal{I}) = \mathcal{M}(\mathcal{I}_1) = \mathcal{M}(\mathcal{I}_2)$, in Corollary 4.4. The proof follows directly from Theorem 3.1.

Corollary 4.4. If $\mathcal{M}(\mathcal{I}) = \mathcal{M}(\mathcal{I}_1) = \mathcal{M}(\mathcal{I}_2)$, then equation (4.19) is satisfied for all $p \in \mathcal{M}(\mathcal{I})$ iff

$$\mathscr{I}_2 = \lambda \mathscr{I}_1 \quad \text{for some} \quad \lambda > 0$$

and

$$\mathscr{C}\begin{pmatrix} X_{12}'X_{11} \\ X_{22}'X_{21} \end{pmatrix} \subset \mathscr{C}\begin{pmatrix} X_{12}'X_{12} \\ X_{22}'X_{22} \end{pmatrix}$$

Examples where these results are valid may be easily obtained by considering two block designs with the same incidence matrix, and having isomorphic designs in their common blocks.

Finally, we remark that in many cases superadditivity of information matrices is due to the fact that some observations may not be used in the BLUEs based on individual experiments due to confounding, but may be released and utilized for estimating linear functions of θ_1 when the data from the experiments are combined. This is clear from the examples considered in this and the previous section.

5. CONCLUDING REMARKS.

In this report we have established the superadditivity of information matrices and explored conditions for additivity. We have also found conditions when BLUEs for the combined experiment can be computed simply from the BLUEs of the individual experiments. The conditions provide geometrical and statistical insight into problems associated with combining experiments. The results are for a general linear model. In any particular setting they may have to be translated to a more readily verifiable form. As an

example, we may refer to Corollary 3.1 where the conditions for the equality $\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2$ have been expressed in a very elementary form for the block design setup.

We hope that the results in this report may be profitably utilized in further research. As an illustration we can consider a problem of extending an experiment optimally. Suppose an experimenter has conducted an experiement for comparing treatments in b block $B_1,\ldots,B_b$ of size k each. Then suppose he is given funds to conduct another experiment of the same nature to improve on his findings. Suppose he has the choice of either using new blocks $B_{b+1},\ldots,B_{2b}$, or using the old blocks again. Suppose the experiment is such that the old blocks have no residual effect from the earlier experiment, i.e., the model remains the same. The question is, how should the experimenter extend the experiment to get the best results? What should be his new design?

Let $\mathcal{I}_1$ be the information matrix of the old design, $\mathcal{I}_2$ the information matrix of the extension and $\mathcal{I}$ the information matrix of the combined experiment. If the experimenter uses new blocks, then surely $\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2$. In fact, if he uses some or all of the old blocks, then also $\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2$ as long as the designs in the common blocks are isomorphic. Thus, he should use the same blocks and try not to have the new design isomorphic to the old one. Obviously, if the old design can be extended to an optimal design in b blocks of size $2k$ each

(for example, a Balanced Block Design), then the information in the combined experiment has been "maximized". Otherwise, one suspects that in general the extension should be as far as possible from designs which are isomorphic to the old design.

Now suppose the experimenter has only two choices: to repeat the same experiment (same design) in new blocks or in the old blocks. Clearly the information matrix is the same in both cases, viz $\mathcal{I}_1 + \mathcal{I}_2$, where $\mathcal{I}_1$ and $\mathcal{I}_2$ have their usual meanings. But if he uses old blocks he saves a lot of degrees of freedom which would be otherwise used for the new block effects. He can use them to test the validity of his model. Alternately, he can use this to get a better estimate of σ^2, the measurement error. In this case, though he is estimating the estimable treatment parameters with the same precision as he would have if he uses new blocks, his estimate of the variance of estimate improves considerably. We may remark that unfortunately the precision of estimators of σ^2 as has been suggested by Fisher (1971) has been largely ignored in the literature of optimal design of experiments.

REFERENCES

1. Fisher, R.A. (1971). The Design of Experiment, Hafner Publishing Co., New York (first published in 1935).

2. Rao, C.R. (1965). Linear Statistical Inference And Its Applications. Second Edition (1973). John Wiley, New York.

3. Rao, C.R. and Mitra, S.K. (1971). Generalized Inverse of Matrices and Its Applications. John Wiley, New York.